

## **2017 Invited Workshops**

### **Title: Teaching Genomic Medicine: A Train-the-Trainer Workshop**

**Description:** There is a clear need to educate clinical trainees in genomic medicine. Traditional lecture formats, however, do not allow for the interactive use of online genomics tools or accommodate today's "millennial learner." This workshop is geared towards genetics professionals who teach or plan to teach genomics to clinical trainees. Participants will gain hands-on experience with a field-tested genomics curriculum that utilizes a team-based learning and flipped classroom approach. In the first portion of the session, participants will be introduced to these teaching methods and the overall structure of the curriculum, which addresses both somatic and germline genomic testing, patient communication and ethical issues. In the second portion, participants will work in teams as "students" on a portion of the curriculum to better understand curricular content and the dynamics of team-based learning. In the final segment, there will be a panel discussion of tips to implement this novel approach to genomics education and how the curriculum has been adapted to various medical specialties. There will also be an opportunity for participants to share their own ideas and experiences. The overall goal is for participants to be able to implement this novel approach to genomics education at their home institutions.

### **Title: Navigating The ENCODE Encyclopedia: Exploring candidate regulatory elements, linked genes, and genetic variation with SCREEN**

**Description:** The Encyclopedia of DNA Elements (ENCODE) Consortium has generated thousands of genomic datasets with the goal of annotating non-coding regions of the genome. In order for the research community to easily use and interpret these datasets, the ENCODE Consortium has integrated multiple datasets to create a collection of genomic annotations termed the ENCODE Encyclopedia. In this workshop, members of the ENCODE Data Analysis Center will introduce the ENCODE Encyclopedia, focusing primarily on the database and visualization tool SCREEN: Search Candidate Regulatory Elements by ENCODE. SCREEN is an online tool that enables users to explore candidate REs across hundreds of cell and tissues types and filter regions by various facets. Participants will use SCREEN to work through hands-on examples focusing on comparing enhancer-like elements across tissues, predicting enhancer-gene interactions, and annotating genetic variants. Exercises will demonstrate how to access and download supporting data, as well as visualize regions of interest using dynamic graphs and external genome browsers. Participants will also learn how to use SCREEN to analyze their own datasets which they are encouraged to bring. Participants will learn how to compare gene expression profiles, intersect regions, and correlate signal for datasets of interest with the ENCODE Encyclopedia.

### **Title: Accessing the breadth of data in Ensembl: a worked clinical example**

**Description:** The Ensembl genome browser ([www.ensembl.org](http://www.ensembl.org)) provides visualisation and analysis of integrated genomic data, including genes, variants, comparative genomics and gene regulation, for over 70 species. This workshop provides guidance for wet-lab scientists and clinicians who have not yet discovered the power and depth of this resource. A brief introduction to Ensembl will be followed by hands-on demonstrations and exercises that describe and demonstrate an adaptable workflow for exploring and exporting genomic data. The workflow includes: 1) use BioMart to export a list of genes falling within a clinically relevant genomic location; 2) from the gene list exported from BioMart, explore a single gene in Ensembl, exporting its sequence and identifying variants, homologues and GO terms; 3) work through a deep dive exploration of a variant that falls within the gene of interest, finding functional consequences, population frequencies from the 1000 Genomes project, population transcript haplotypes and phenotypes; and 4) explore GTEX eQTL data to discover relevant variants in the region surrounding the gene of interest, that affect its expression in various tissues. Workshop materials, including slides, demonstration screenshots, exercises and solutions will be made available before the workshop and will remain permanently online at our training portal: <https://training.ensembl.org>.

**Title: Overview, interpretation and use of the GTEx resources: eQTLs and gene expression**

**Description:** Identifying genetic variation regulating gene expression remains challenging, since regulatory motifs are found both local and distal to their target gene, and are difficult to identify based on sequence alone. Moreover, gene expression and its regulation vary greatly among tissues and cell types, many of which are difficult to sample. The Genotype-Tissue Expression (GTEx) project provides an unprecedented breadth of transcriptomic (RNA-seq) and genetic (WGS and WES) data from >50 non-diseased human tissues across ~960 donors. This interactive workshop will showcase the GTEx data resource, and demonstrate how to access, interpret, and visualize the various types of data produced, using downloadable files, the GTEx Portal, and API. Given the complexity of interpreting genetic associations with gene expression, we will review computing expression quantitative trait loci (eQTLs), detecting independent signals at a given locus, and narrowing down a high confident set of causal variants per eQTL region. We will demonstrate how to run GTEx analysis pipelines on the FireCloud platform to enable integration of users' own datasets with this resource. Through case studies (e.g., using eQTLs to interpret genetic associations with complex traits; interpreting RNA-seq expression data across tissues), we will also discuss best practices and caveats for using GTEx data.

**Title: Effective Communication of Genetic Research**

**Description:** The use of highly technical language is critical for genetic researchers to communicate precisely with each other about science. When it comes to communication of the same concepts outside the community – even to scientists in other fields – this language becomes easily misinterpreted. These misunderstandings are a hindrance to our ability to make our science enticing to funding agencies, to work effectively with collaborators outside our specific field, to recruit volunteers into our studies, to be accurately interpreted by writers, and to communicate with patients and their families. This workshop is designed to foster effective communication skills about our research. Active group exercises between workshop participants and panelists will be used to identify areas in which miscommunication occurs and best practices for communicating about science, including ways to sell your research without overstating it. Participants with a clinical focus will gain skills simplifying complex genetic information while maintaining precision.

**Title: UCSC Genome Browser - new data, new features**

**Description:** The UCSC Genome Browser is a widely used visualization tool for access to genomic data of many kinds. Continuing development and new datasets make the Browser more useful each year. The ability to load your own RNA-seq, whole-exome and whole-genome data into the Browser allows co-visualization with known benign and pathogenic variants for evaluation. The Variant Annotation Integrator predicts biochemical consequences of variants. The Browser now supports HGVS nomenclature for navigation. New data sets display pre-computed evaluations of the optimal locations to use CRISPR technology to modify the genome, and GTEx tissue-specific gene expression data are now visible in the Browser graphic. Finally, multi-region display mode offers visualization of discontinuous sections of the genome, including multiple regions on any chromosome side-by-side, making the display of RNA-seq and all other annotation data across a pathway much more intuitive in a single view. Facility with the Browser to access an integrated view of genomic data will enable researchers and clinicians alike to make the most of a wide range of whole-genome annotations available all in one place. Attendees should have some familiarity with the Genome Browser and bring fully charged laptops to participate.

**Title: Classifying and interpreting germline and somatic variants in your large cohort study with the CRAVAT tool suite**

**Description:** Variant classification is a critical issue for research and clinical labs involved in germline and somatic sequencing, particularly those looking at high-risk patient populations. CRAVAT is a free-to-use toolkit for interpreting genomic variation, developed at Johns Hopkins, funded by NIH/NCI, and continually improved to reflect input from our base of 1000s of clinical and experimental users from 188 countries across six continents. As such, CRAVAT delivers consensus standards and innovative protocols for variant interpretation in a highly intuitive, visually stunning, and interactive environment. CRAVAT is fast, results are explored through a secure interactive dashboard with publication-ready plots, verbose annotation, and 3D protein structural mapping, all of which can be accomplished via our web portal, in the cloud, or run locally. Our workshop is a hands-on presentation that uses publicly available germline and somatic variation data to cover job submission, scoring of variants, dynamic filtering to identify interesting variants, and use of visualization and drill-down panels to gain insight into the functional impact of important variants. Participants will learn how to leverage CRAVAT to optimize the cycle of variant discovery and functional interpretation.

**Title: CRISPR-Cas9 Genome Editing Bootcamp**

**Description:** Genome-editing tools, particularly CRISPR-Cas9, are becoming integral to the investigation of human genetics. This hands-on workshop is intended for individuals interested in learning practical aspects of CRISPR-Cas9-based technologies. The workshop follows a flipped-classroom, active-learning model, with participants viewing an online video before the meeting. During the workshop, participants work in teams on virtual experiments in which genome-edited knockout/knock-in mice and cell lines are generated (requiring participants to bring their own WiFi-capable laptops). Participants will learn online tools to design CRISPR-Cas9 reagents, common pitfalls to avoid, and data interpretation from genome-editing experiments, with guidance from experienced AHA functional genomics and translational biology instructors.

## 2016 Invited Workshops

**Title:** ASHG Interactive Workshop: Computer Aided Dysmorphology Training

**Description:** Dysmorphology training is a real challenge. The rarer the syndrome, the lower the probability of a medical professional seeing that syndrome more than once. Computer-aided dysmorphology analysis enables the clinical geneticist to benefit from the cumulative knowledge of geneticists worldwide. This tool will allow workshop participant to receive a hand-on learning experience and test their dysmorphology knowledge and skills. This workshop will include a 15 min overview of best practices in using computer-aided dysmorphology analysis followed by a 15 minute analysis of 2 cases using this tool. This will be followed by 45 min. practical training using specific exercises /challenges on each participant's laptop or tablet computer. Participants will be able to use references and interconnected dysmorphology databases in solving these exercises /challenges. Q&A will be allowed throughout the practical exercises.

**Title:** ASHG Interactive Workshop: Introduction to Research Narratives with GenePattern Notebooks

**Description:** As the availability of genetic and genomic data and their associated software tools continues to increase, the need has become urgent for an analysis environment that supports the entire "idea to dissemination" cycle of an integrative genomics analysis. GenePattern Notebook unites these phases of in silico research - experiment design, collaborative analysis, and publication - into a single interface. GenePattern Notebook supports the full research life cycle: the same environment that enables incremental, step-by-step research also provides a finished document, text and embedded analysis, that can serve as supplementary material for a publication. GenePattern Notebook, based on the GenePattern environment for integrative genomics and the Jupyter Notebook system, provides both a familiar notebook format that gives non-programming investigators easy access to multi-step analyses, and also a flexible and powerful analysis and visualization environment in which experienced bioinformaticians can mix GenePattern modules with code in Python, R, and other supported languages. Participants will learn how to use GenePattern Notebook to create, share, and disseminate executable research documents that interleave their analyses with descriptive text and graphics, all available with no requirement for programming expertise.

**Title:** ASHG Interactive Workshop: Open Science: Quality Assurance and Analysis of ChIP-seq Data using the ENCODE Uniform Processing Pipeline.

**Description:** Cooperative consortia such as ENCODE, Epigenomics Roadmap, BLUEPRINT, and others have generated ChIP-seq data from diverse cell types and tissues. These experiments establish a baseline measurement of the distribution of histone modifications, transcription factors, and chromatin regulators in undifferentiated, differentiated, and transformed cells. The value of this baseline measurement for comparative analysis depends on uniform processing. Within the ENCODE consortium, ChIP-seq experiment protocols, quality metrics, and data analysis methodologies have been standardized so that results from different laboratories can be compared directly. At this workshop, attendees will run the ENCODE ChIP-seq pipeline in the cloud, shepherding the raw sequence data (fastq) from a ChIP-seq experiment, mapping these sequences to the human reference, calling peaks and signal output and visualizing it at the UCSC genome browser. Furthermore, ENCODE has implemented the data analysis pipelines in a form that can be replicated by anyone on ENCODE data, data from other projects, data from the literature and public databases, as well as data from new ChIP-seq experiments. All of the code is open-source, and the analysis pipelines are deployed on a cloud computing platform with an interface that is entirely web-based so that no IT or HPC infrastructure is necessary. Users are guaranteed that the quality metrics and results from their analyses are directly comparable to all ENCODE data and any other data processed through the uniform ENCODE pipeline.

**Title:** ASHG Interactive Workshop: Teaching Genomic Medicine: A Train-the-Trainer Workshop

**Description:** There is a clear need to educate clinical trainees in genomic medicine. Traditional lecture formats, however, do not allow for the interactive use of online genomics tools or accommodate today's "millennial learner." This workshop is geared towards genetics professionals who teach or plan to teach genomics to clinical trainees. Participants will gain hands-on experience with a field-tested genomics curriculum that utilizes a team-based learning and flipped classroom approach (<http://www.pathologylearning.org/trig/resources>, <http://www.ncbi.nlm.nih.gov/pubmed/24678680>). In the first portion of the session, participants will be introduced to

these teaching methods and the overall structure of the curriculum, which addresses both somatic and germline genomic testing as well as patient communication and ethical issues. In the second portion, participants will work in teams as "students" on a portion of the curriculum to better understand curricular content and the dynamics of team-based learning. In the final segment, there will be a panel discussion of tips to implement this novel approach to genomics education and how the curriculum has been adapted to various medical specialties. There will also be an opportunity for participants to share their own ideas and experiences. The overall goal is for participants to be able to implement this novel approach to genomics education at their home institutions.

**Title:** ASHG Interactive Workshop: Introduction to the Integrative Genomics Viewer (IGV)

**Description:** This workshop provides an introduction to the Integrative Genomics Viewer (IGV), a widely used desktop application for interactive visual exploration of a wide range of genomic data types, including sequence alignments, genomic annotations, copy-number, gene expression, and clinical data. In particular, participants will learn how to use IGV to view next generation sequence (NGS) data, including: (1) the basics of using the IGV application, (2) the different options IGV provides for the visualization of NGS alignment data, and (3) how IGV supports visual inspection of the data, including variant call validation, and interpretation of insert sizes and pair orientation in paired-end sequencing data. Participants will leave the workshop with the ability to use IGV to visualize NGS data for discovery as well as validation. The workshop includes both lectures and hands-on exercises.

**Title:** ASHG Interactive Workshop: Bringing your Classroom into the Genomic Era

**Description:** Although the core tenets of inheritance do not change, the Human Genome Project and its offshoots have greatly accelerated the speed at which genomic technologies have evolved. In this genomic era, genetics curricula are at risk of rapidly becoming out-of-date. In parallel, there has been increased emphasis on introducing opportunities for active learning into higher education. This workshop is designed to meet these challenges and blends the goals of modernizing genetics content and teaching approaches. The workshop is geared towards genetics educators at all levels, from undergraduate to post-graduate. Participants will identify learning objectives that are commonly not being met in their curricula and will work together to develop a lesson plan to address these objectives. Participants will first discuss cutting edge genomics concepts and emerging active learning pedagogical approaches before selecting an area on which to focus. Subsequently, participants will work in groups to design a session that will address these focus areas, including interprofessional education, whole genome and exome sequencing, and cancer genomics. We aim to have the participants walk out with a deliverable case that can be implemented in their home institutions.

**Title:** ASHG Interactive Workshop: New Resources at the UCSC Genome Browser for the Display and Interpretation of Sequencing Data

**Description:** The UCSC Genome Browser is a widely used platform for access to genomic data provided by laboratories around the world and for display of user data alongside it. New tools are constantly being developed by our team, and even regular users are often not aware of the full potential of the Browser suite of tools. In this intermediate-level workshop, you will learn how to load your own data, including whole-exome and RNA-seq, into the Browser; how to intersect it with other data (yours or ours) to find regions where the signals reinforce (or do not) and how to export lists of genes and/or SNPs from those regions. The new Data Integrator will play a key role in the analysis, as will the new exon-only display mode, which allows condensed viewing for visualization of whole-exome sequencing results. Finally, participants will learn how to use the Genome Browser-in-a-Box, a virtual-machine version of the Browser that provides enhanced data security and privacy by operating within an institutional firewall. Previous familiarity with the Genome Browser is recommended, but not required.

**Title:** ASHG Interactive Workshop: 3D Genome Organization and Chromatin Interaction

**Description:** The three dimensional (3D) organization of mammalian genomes is tightly linked to gene regulation, as it reveals the physical interactions between distal regulatory elements and their target genes, and further orchestrate spatial- and temporal-specific gene expression. Several recent high-throughput technologies based on Chromatin

Conformation Capture (3C) have emerged (such as 4C, 5C, Hi-C and ChIA-PET) and given us an unprecedented opportunity to study the higher-order genome organization. However, genome-wide chromatin interaction data analysis and visualization are complicated and even in tissue/cell types where the data are published, it is still difficult for the scientific community to visualize and use them for their own research. This interactive workshop will showcase three online tools to illustrate how to query gene expression, candidate cis-elements (such as promoters and enhancers), chromatin interaction loops, and 3D genome organization. All the presenters are experienced investigators from the ENCODE, Roadmap/Epigenomics and newly founded 4D Nucleosome consortia. This workshop will integrate the data generated through the aforementioned large projects, with a focus gene regulation through chromatin interaction. We will use case studies to show the audience how disease-related genetic variants can lead to altered gene expression and eventually disease pathogenesis.

**Title:** ASHG Interactive Workshop: Variant Discovery with GATK 4

**Description:** The Genome Analysis Toolkit or GATK, developed at the Broad Institute, is one of the most widely used software packages for variant discovery in whole genome and exome data. This workshop will introduce the core "GATK Best Practices", a complete solution for variant discovery that goes from data preparation to variant calling, genotyping, filtering and evaluation. We will focus the hands-on exercises on the latest methodological innovations in the upcoming GATK version 4 (scheduled for release in March 2016) that will enable researchers to obtain best in class results from their data with cost-effective performance. This hands-on session is designed for attendees to follow along and run the exercises on their own laptop. All necessary materials and installation instructions will be provided at time of registration, and attendees will be required to have completed installation prior to the workshop. Attendees will be encouraged to participate actively through mini-quizzes. A complete package of course materials including example data, detailed exercise walkthroughs and supporting documentation will be made available for download.